

Chapitre 1 :Rappels et compléments d'analyse matricielle

Chapitre 2 : fondements du calcul scientifique

Chapitre 3 : Méthodes directes de résolution des systèmes linéaires

■ Chapitre 4 : Méthodes itératives de résolution des systèmes linéaires

Chapitre 1 :Rappels et compléments d'analyse matricielle

Chapitre 2 : fondements du calcul scientifique

Chapitre 3 : Méthodes directes de résolution des systèmes linéaires

■ Chapitre 4 : Méthodes itératives de résolution des systèmes linéaires

2.1 Sources d'erreurs dans un modèle numérique

2.2 Evaluation de l'erreur

2.3 Propagation d'erreur

2.4 Représentation des nombres en machine

2.5 Troncature et arrondi d'un nombre

2.6 Arithmétique flottante

2.7 Conditionnement

2.1 Sources d'erreurs dans un modèle numérique

Exemple introductif

On cherche à évaluer de manière numérique l'exponentielle $z = \exp(x)$.

⇒ Trouver z^* approchant numériquement z .

Une méthode consiste à utiliser les séries de Taylor

$$\forall x \in \mathbb{R} \quad \exp(x) = \sum_{k=0}^{+\infty} \frac{x^k}{k!}$$

En considérant des ressources de calcul limitées, on en est réduit à déterminer

$$z^* = \sum_{k=0}^n \frac{x^k}{k!}$$

⇒ Pour le calcul de l'exponentielle, on commet donc une erreur.

❏ Erreurs de modélisation

- **Erreurs de modèle** : ces erreurs sont dues au fait que les modèles mathématiques sont plus ou moins idéalisés, ce qui donne lieu à plusieurs erreurs.

Exemple l'erreur du modèle du pendule qui ne tient pas en considération la force de friction.

- **Erreurs de mesure** : ces erreurs sont dues à la présence dans le modèle mathématique de paramètres numériques dont les valeurs ne peuvent être observées ou déterminées qu'approximativement suite à des mesures expérimentales.

Telles sont toutes les constantes physiques, comme, par exemple, la longueur l dans le modèle du pendule.

❏ Erreurs numériques

- **Erreurs d'approximation ou de troncature** : ces sont les erreurs associées aux processus infinis en analyse mathématique.
Exp: les séries numériques.
- **Erreurs d'arrondi** : ce sont les erreurs associées à la représentation des nombres en machine: un ordinateur ne peut prendre en considération qu'un nombre fini de chiffres.
- **Erreurs de propagation et génération** : ces sont les erreurs qui apparaissent dans le résultat d'une opération comme conséquence des erreurs des opérandes.

Dans ce qui suit nous allons nous intéresser aux deux derniers types d'erreur.

2.2 Evaluation de l'erreur

Soient x une valeur exacte et x^* une valeur approchée, on définit:

□ **Erreur absolue** $E = |x - x^*|$, sert à déterminer la précision de la valeur approchée x^* par rapport à la valeur exacte x .

Exemple:

Pour la valeur exacte $x = 2/3$, la valeur approchée $x_1^* = 0,666667$ est mille fois plus précise que la valeur approchée $x_2^* = 0,667$.

En effet, nous avons

$$E_1 = |x - x_1^*| = |2/3 - 0,666667| = \frac{1}{3} \cdot 10^{-6}$$

$$E_2 = |x - x_2^*| = |2/3 - 0,667| = \frac{1}{3} \cdot 10^{-3}$$

❑ **Erreur relative** $E_r = \frac{|x - x^*|}{|x|}$, sert à comparer la précision de différentes valeurs approchées x^* ; y^* relativement à différentes valeurs exactes x ; y .

En multipliant par 100 on obtient l'erreur relative en pourcentage.

Exemple : Pour les valeurs exactes $x = 2/3$ et $y = 1/15$, on considère les valeurs approchées respectives $x^* = 0,67$ et $y^* = 0,07$.

Les erreurs absolues sont: $E_1 = |x - x^*| = |2/3 - 0,67| = \frac{1}{3}10^{-2}$

$$E_2 = |y - y^*| = |1/15 - 0,07| = \frac{1}{3}10^{-2}$$

Les erreurs relatives sont : $E_{r1} = E_1/|x| = 0,5\%$

$$E_{r2} = E_2/|y| = 5\%$$

Ainsi, bien que les erreurs absolues soient égales, x^* est une approximation dix fois plus précise pour x que ne l'est y^* pour y .

□ En pratique on ne connaît pas x et E .

On définit alors un majorant de l'erreur absolue Δx d'une valeur approchée x^* par :

$$E = |x - x^*| \leq \Delta x \Leftrightarrow -\Delta x \leq x - x^* \leq \Delta x$$
$$\Leftrightarrow x^* - \Delta x \leq x \leq x^* + \Delta x$$

tel que x est un nombre réel positif.

□ On définit un majorant de l'erreur relative δx d'une valeur approchée x^* par

$$\delta x = \frac{\Delta x}{|x^*|}$$

Exemple

1) Une surface est donné par $x = 60\text{m}^2 \pm 2\%$.

- la valeur approchée de la valeur exacte x est $x^* = 60\text{m}^2$
- L'erreur relative à x^* est $\delta x = 0.02$
- L'erreur absolue $\Delta x = x^* \delta x = 60 \times 0.02 = 1.2\text{m}^2$
- la surface exacte est $x \in [x^* - \Delta x, x^* + \Delta x] = [58.8, 61.2]$.

2) Un volume est donné par $x = 512.4\text{m}^3 \pm 1.5$

- la valeur approchée de la valeur exacte x est $x^* = 512.4\text{m}^3$
- L'erreur absolue à x^* est $\Delta x = 1.5$
- L'erreur relative est $\delta x = \frac{\Delta x}{x^*} = \frac{1.5}{512.4} = 0.3\%$

2.3 Propagation d'erreur

Soient x et y deux valeurs exactes, x^* et y^* leur approximation respective

Opération	Erreur absolue	Erreur relative
$x^* + y^*$	$\Delta x + \Delta y$	$\frac{\Delta x + \Delta y}{ x^* + y^* }$
$x^* - y^*$	$\Delta x + \Delta y$	$\frac{\Delta x + \Delta y}{ x^* - y^* }$
$x^* \times y^*$	$x^* \Delta y + y^* \Delta x$	$\frac{\Delta x}{x^*} + \frac{\Delta y}{y^*}$
x^* / y^*	$\frac{(x^* \Delta y + y^* \Delta x)}{(y^*)^2}$	$\frac{\Delta x}{x^*} + \frac{\Delta y}{y^*}$

2.4 Représentation des nombres

- Un nombre est connu numériquement si l'on dispose de son développement décimal

- **un entier**: développement décimal limité

$$9325 = 9 \cdot 10^3 + 3 \cdot 10^2 + 2 \cdot 10^1 + 5 \cdot 10^0$$

- **nombre rationnel**: développement décimal illimité et périodique

$$1/7 = 0,142857142857142857.....$$

$$= 1 \cdot 10^{-1} + 4 \cdot 10^{-2} + 2 \cdot 10^{-3} + 8 \cdot 10^{-4} + 5 \cdot 10^{-5} + 7 \cdot 10^{-6} +$$

- **nombre réel**: développement décimal illimité

$$0,00123456789..... = 1 \cdot 10^{-3} + 2 \cdot 10^{-4} + 3 \cdot 10^{-5} + 4 \cdot 10^{-6} +$$

- Un ordinateur ne peut représenter qu'un sous-ensemble fini de l'ensemble des nombres réels → toute opération d'un ordinateur est entachée par des erreurs d'arrondi.

☐ Chiffres significatifs

- Un chiffre significatif d'un nb approché est le seul chiffre qu'on doit garder càd
 - tout chiffre dans sa représentation décimale $\neq 0$
 - un zéro s'il se trouve entre deux chiffres significatifs ou s'il constitue un chiffre conservé

Exemple: $x=0,00301045$. l'approximation à 6 décimales :

$$x^* = 0,003010 = 3 \cdot 10^{-3} + 0 \cdot 10^{-4} + 1 \cdot 10^{-5} + 0 \cdot 10^{-6}$$

- **Comment déterminer le nb de chiffres significatifs d'une valeur?**

Les chiffres significatifs correspondent à l'ensemble des chiffres apparaissant à partir du 1^{er} chiffre $\neq 0$ allant de gauche vers la droite

x^*	3000,05	7300	0,0073	30	3,5889	0,00009	023,2
Nb chif sig	6	4	2	2	5	1	3

Attention !! La puissance de dix n'intervient pas dans le décompte

Exemple : la valeur de la vitesse de la lumière mesurée à $3,00 \times 10^8$ m/s possède trois chiffres significatifs : le 3 et les deux 0.

Les chiffres significatifs

102,0 : 4 chiffres significatifs

1 2 3 4

Les zéros situés à droite sont significatifs
(ils indiquent la précision correcte : dans cet exemple au dixième près)

~~00~~10700 : 5 chiffres significatifs

1 2 3 4 5

On compte à partir du premier chiffre à gauche non nul
Les zéros situés à gauche ne sont PAS significatifs
(ils disparaissent si on choisit une unité plus petite et un changement d'unité ne change pas la précision)

102,0 $\times 10^5$: 4 chiffres significatifs

1 2 3 4

La puissance de dix n'intervient pas dans le décompte

- Les chiffres **significatifs exacts** correspondent à tous les chiffres connus avec certitude. $X = 1/3 = 0,33333333\dots$
 $x^* = 0,33334$
- si ***m*** est l'entier tel que $\Delta x \leq 0.5 \cdot 10^m$,
 alors le chiffre correspondant à la ***m***^{ème} puissance de 10 **de x^*** et tous ceux à sa gauche sont des chiffres significatifs exacts.

Exemple:

$$x = \pi = 3.14159265358\mathbf{979} \dots, \quad \tilde{x} = 22/7 = 3.142857\dots$$

$$\Delta x = 0.00126 = 0,126 \cdot 10^{-2} < 0.5 \cdot 10^{-2}$$

$$(\text{ si par exemple } \Delta x = 0.00726 =$$

$$= 0.726 \cdot 10^{-2}$$

$$= 0,0726 \cdot 10^{-1} < 0.5 \cdot 10^{-1}$$

le chiffre des centièmes est significatif et au total on a trois chiffres significatifs exacts.

- Si un nbre approximatif possède n chiffres significatifs, alors l'erreur relative est $< 5 \cdot 10^{-n}$

Exemple: soit $x^* = 34.152$, et on sait que les 5 chiffres sont significatifs exacts. L'erreur relative dans ce nombre est donc $< 5 \cdot 10^{-5}$.

- Soit x^* une approximation de x . Si l'erreur relative dans x^* est $\leq 0.5 \cdot 10^{-n}$, alors x^* possède au moins n chiffres significatifs exacts.

Exemple: si on sait que l'erreur relative dans 345.24 est $< 0.5 \cdot 10^{-4}$, on peut conclure qu'au moins les chiffres 3, 4, 5, et 2 sont exacts. Le chiffre 4 reste incertain.

□ Représentation en virgule flottante

Ecriture de nombres réels en machines. Soit $x \in \mathbb{R}$, sa représentation en *virgule flottante* est

$$\pm m b^e = \text{signe} \times \text{mantisse} \times b^{\text{exposant}}$$

b = la base du système numérique utilisé (base 2,8,16,10,...)

m = la mantisse $1/b \leq m \leq 1$, suite d'entier tels que

$m = 0, a_1 a_2 \dots a_n$ $0 \leq a_i \leq b-1$, $i=1,2,\dots$

pour garantir l'unicité de la représentation on impose que le premier chiffre après la virgule est $\neq 0$

e =l'exposant(un nombre entier relatif)

Exemple: pour $b=10$

$$x = 3,4 \quad \rightarrow \quad fl(x) = 3,4 \cdot 10^{-1} \cdot 10^1 = 0,34 \cdot 10^1$$

$$x = 0,0038 \rightarrow fl(x) = 0,0038 \cdot 10^2 \cdot 10^{-2} = 0,38 \cdot 10^{-2}$$

2.5 Troncature et arrondi d'un nombre

Soient x un réel non nul et $fl(x)$ sa représentation en virgule flottante à n chiffres

- **La troncature** consiste à retrancher les chiffres à partir de la position $n+1$. L'erreur relative maximale est de l'ordre de 10^{-n}

Exemple: $n=4$

$$fl(0,027457) = fl(0,2745\cancel{7} \cdot 10^{-1}) = 0,2745 \cdot 10^{-1}$$

- **Arrondi** consiste à ajouter un 5 au $(n+1)$ -ième chiffre avant d'effectuer la troncature. L'erreur relative maximale est de l'ordre de $5 \cdot 10^{-(n+1)}$

Exemple: $n=4$

$$fl(0,027457) = fl(0,2745\color{red}{7} \cdot 10^{-1}) = 0,2746 \cdot 10^{-1}$$

$$fl(0,027453) = fl(0,2745\color{red}{3} \cdot 10^{-1}) = 0,2745 \cdot 10^{-1}$$

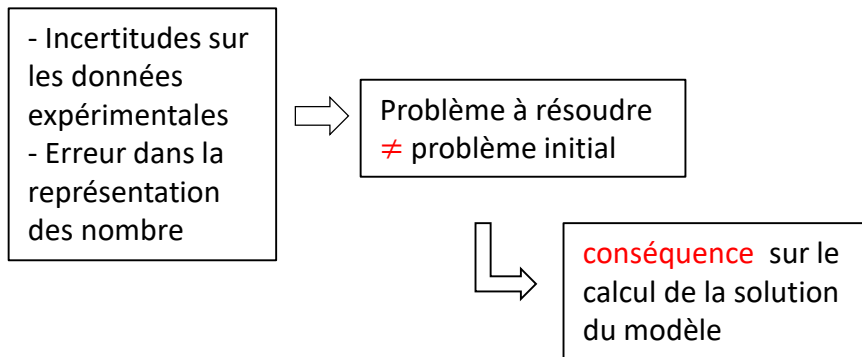
2.6 Arithmétique flottante

- Pour effectuer une opération sur deux nombres réels, on effectue l'opération sur leurs représentations flottantes et on prend ensuite la représentation flottante du résultat:

$$op \in \{+, -, \div, \times, \sqrt{}\}, \quad x \text{ op } y = fl(fl(x) \text{ op } fl(y))$$

- Chaque opération intermédiaire dans un calcul introduit une nouvelle erreur d'arrondi ou de troncature.
- On ne peut effectuer l'addition ou la soustraction que si les exposants sont les mêmes. On transforme le plus petit exposant et on ne respecte plus la règle voulant que le premier chiffre de la mantisse ne soit pas nul.
- Deux expressions algébriquement équivalentes peuvent fournir des résultats différents et que l'ordre des opérations peut changer les résultats.

2.7 Conditionnement



- C'est la notion de **conditionnement** d'un problème qui est une mesure de la difficulté de calcul numérique du problème.
- On dit qu'un problème est bien (ou mal) conditionné, si une petite variation des données entraîne une petite (une grande) variation sur les résultats.